Water Research 106 (2016) 383–393

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Global resilience analysis of water distribution systems

Kegong Diao [a, b, *], Chris Sweetapple [a, **], Raziyeh Farmani [a], Guangtao Fu [a], Sarah Ward [a], David Butler [a]

[a] Centre for Water Systems, University of Exeter, North Park Rd, Exeter, EX4 4QF, UK
[b] Faculty of Technology, De Montfort University, Mill Lane, Leicester, LE2 7DR, UK

ARTICLE INFO

Article history:
Received 15 December 2015
Received in revised form
13 September 2016
Accepted 3 October 2016
Available online 4 October 2016

Keywords:
Excess demand
Failure mode
Global resilience analysis
Pipe failure
Substance intrusion
Water distribution system

ABSTRACT

Evaluating and enhancing resilience in water infrastructure is a crucial step towards more sustainable urban water management. As a prerequisite to enhancing resilience, a detailed understanding is required of the inherent resilience of the underlying system. Differing from traditional risk analysis, here we propose a global resilience analysis (GRA) approach that shifts the objective from analysing multiple and unknown threats to analysing the more identifiable and measurable system responses to extreme conditions, i.e. potential failure modes. GRA aims to evaluate a system's resilience to a possible failure mode regardless of the causal threat(s) (known or unknown, external or internal). The method is applied to test the resilience of four water distribution systems (WDSs) with various features to three typical failure modes (pipe failure, excess demand, and substance intrusion). The study reveals GRA provides an overview of a water system's resilience to various failure modes. For each failure mode, it identifies the range of corresponding failure impacts and reveals extreme scenarios (e.g. the complete loss of water supply with only 5% pipe failure, or still meeting 80% of demand despite over 70% of pipes failing). GRA also reveals that increased resilience to one failure mode may decrease resilience to another and increasing system capacity may delay the system's recovery in some situations. It is also shown that selecting an appropriate level of detail for hydraulic models is of great importance in resilience analysis. The method can be used as a comprehensive diagnostic framework to evaluate a range of interventions for improving system resilience in future studies.

1. Introduction

There is an emerging realisation that building resilience is an important component of enhancing the sustainability of many systems, including water systems (Ahern, 2011; Pickett et al., 2014). Engineering resilience can be broadly characterised in two related but distinct ways: attribute-based and performance-based. The former typically concerns the system as a whole and could be considered as a set of design principles, such as the degree of interconnectedness or duplication, which enables the system to respond appropriately to any threat. The latter refers to the agreed performance of the system (or part of the system) in responding to a particular threat. It is therefore typically prescriptive (*i.e.* standard-based) and refers to an operational goal (Butler et al., 2014). The degree to which the various attributes of a system build the standard of performance required is still a matter of ongoing research but this requires a detailed understanding of whole-system resilience.

In this paper, emphasis is placed on understanding the performance of water systems under unexpected or deteriorating conditions (beyond failure) and resilience is defined here as the degree to which the system minimizes level of service failure magnitude and duration over its design life when subject to exceptional conditions (Butler et al., 2014, 2016). This definition is broadly in line with emerging policy and practice (NIAC, 2009; Ofwat, 2012). As a prerequisite to *enhancing* resilience, a detailed understanding is required of the inherent resilience of the underlying system and that is the focus of this paper.

There area number of key challenges to overcome when

* Corresponding author. Faculty of Technology, De Montfort University, Mill Lane, Leicester, LE2 7DR, UK.
** Corresponding author. Centre for Water Systems, University of Exeter, North Park Rd, Exeter, EX4 4QF, UK.
E-mail addresses: dkg630630@gmail.com, kegong.diao@dmu.ac.uk (K. Diao), C.Sweetapple@exeter.ac.uk (C. Sweetapple), R.Farmani@exeter.ac.uk (R. Farmani), G.Fu@exeter.ac.uk (G. Fu), sarah.ward@exeter.ac.uk (S. Ward), D.Butler@exeter.ac.uk (D. Butler).

http://dx.doi.org/10.1016/j.watres.2016.10.011
0043-1354/

evaluating system resilience:

- How to link threat to impact
- How to deal with different threats that produce the same impact
- How to deal with similar threats that produce different impacts
- How to envisage all possible threats that may affect the system
- How to handle unknown threats.

Risk analysis (e.g. all-hazards approach, ASCE Policy Statement 518, 2006; NIAC, 2009; Cabinet Office, 2011) is a typical method to study threat-impact relationships, but is unable to address unknown threats (Hughes and Healy, 2014). A promising way to overcome this problem is to focus on the system, and direct attention to all possible failure modes rather than speculate on all possible threats. Pipe failure in water distribution systems, for example, is a failure mode that may result from various internal (e.g. water hammer) or external threats (e.g. ground movement). Each failure mode may be considered a 'stress' on the system which results in performance 'strains' (Johansson, 2007; Hokstad et al., 2013). Each failure mode encompasses multiple failure scenarios: in the example of the 'pipe failure' failure mode, for instance, failure scenarios would include every potential combination of pipe failures in the network. Each stress may also vary in magnitude – for example from 0% pipe failure to 100% pipe failure. Failure scenarios and their resulting performance strains (measured in terms of impact/level of service), therefore, can span from high probability, routine failures (such as a single pipe failure) to low (or even unknown) probability total failures. As threats with low probability (e.g. Canadian snow storm in 1998, North American blackout 2003, and earthquake followed by a tsunami in Japan 2011) do happen from time to time it is important to include these high stress scenarios when evaluating resilience (Johansson et al., 2013). With regard to events where the probability of a scenario is incalculable, this is precisely where a 'resilience' approach is of greatest benefit – risk cannot be calculated but that does not mean that such occurrences should be ignored, and resilience assessment provides a tool by which they can be considered. Identification of failure modes, failure scenarios and stress magnitudes requires no knowledge of the causal threat or threats. The full range of performance strains resulting from any stress magnitude can, therefore be evaluated even when there are unknown threats.

This type of approach has been adopted as part of the global vulnerability analysis (GVA) method, used to assess the vulnerability of power grids and water distribution systems (Johansson, 2007; Johansson and Henrik, 2010; Johansson et al., 2013; Hokstad et al., 2013). For each failure mode, the method can identify: the range of strain magnitude (minimum and maximum) that can result from a given stress magnitude, the stress magnitude the system can withstand before reaching a certain level of service reduction and the existence of thresholds where a slight increase in stress gives rise to much more severe strains (Johansson, 2007; Hokstad et al., 2013). In a similar way, Gheisi and Naser (2014) explored water distribution system (WDS) reliability in relation to pipe failure mode. In their study, the performance of alternative WDS layout designs was tested under an increasing number of simultaneous failed pipes. The corresponding strain frequency and magnitude was then measured. However, these previous studies did not model the recovery of the system from failure or measure the strain duration. Thus, these methods cannot be directly applied to comprehensively evaluate the resilience of WDSs.

There have been a few attempts to identify the single failure mode scenarios to which the WDS is most vulnerable. Typically, these studies target identification of the minimum stress magnitude that results in the maximum strain, instead of focusing on an overview (e.g. identification of full range of strains at any given stress magnitude). For example, in terms of pipe failure, Berardi et al. (2014) applied an evolutionary algorithm to identify scenarios that have a minimum fraction of failed pipes, yet result in maximum shortage of water supply (e.g. isolation of connections to all water sources). For the same purpose, Kanta (2006), Bristow et al. (2007) and Kanta and Brumbelow (2013) studied pipe failure mode during fire fighting and identified its maximum strain.

This paper proposes a method that builds on GVA, called global resilience analysis (GRA), which is designed to assess the whole-system resilience of engineering systems. GRA is applied in this paper to WDSs and in a complementary paper to urban drainage systems (Mugume et al., 2015). GRA resilience analysis of a water distribution system has not been attempted before. As such, this study addresses potential large magnitude system failures which have not been captured in any previous analysis, in addition to the lower magnitude/higher probability events already studied. Of course, high stress scenarios are highly improbable and it is likely that decision makers will focus on the lower magnitude/higher probability events. However, there are examples of low probability, high consequences events in water distribution systems, e.g. a large number of pipe failures following a number of days below zero temperature or an earthquake. Hence, we cannot be blind to the fact that extreme events can occur, and this study evaluates the maximum theoretical level of service that may be maintained in such instances. There may also be scenarios in which failure of additional components results in no further impact. Therefore, to achieve a comprehensive view, it is important to consider as many scenarios as possible. GRA has been attempted in the context of urban drainage (Mugume et al., 2015) but this study includes specific consideration to both structural and functional failures (see below for definitions of these). Moreover, this study applies an improved failure scenario sampling technique which includes targeted failures in addition to the random failures. This enriches the sample and increases the capture of critical scenarios, particularly those resulting in distinctive high and low strains. In contrast to the methods reviewed above, GRA measures both strain (level of service) magnitude and duration under stresses, allowing a comprehensive picture of system resilience to be built up. Failure scenarios over extended periods (Francis and Bekera, 2014) are modelled for the failure mode. Metrics are based on model simulation results rather than graph theory, as is the case in GVA. This paper describes the GRA method developed and its application to four different case studies. Three different failure modes are considered and resilience curves built, compared and contrasted.

2. Method

2.1. Global resilience analysis

A system's inherent resilience is evaluated by modelling the basic failure modes with increasing stress magnitude and estimating the corresponding strains that arise (Johansson, 2007; Hokstad et al., 2013). The method includes the following steps:

Step 1. Identify the failure mode to be considered (e.g. structural failure, excess demand); this study selects three WDS failure modes, of which details are provided in subsection "Failure modes selected".
Step2. Identify the system stress associated with the failure mode and the way to simulate it (e.g. WDS simulation with excess load at a node for a specified period);
Step 3. Identify the appropriate system strain and how to measure it (e.g. ratio of unsupplied demand to the total demand required in the strain duration);

Step 4. Simulate failure mode strains under increasing stress magnitude (0%–100%of maximum stress). Whilst extreme stress magnitudes of up to 100% may be highly improbable, they are theoretically possible and must, therefore, be included if the full range of potential impacts is to be identified. For any given stress magnitude, an appropriate number of failure scenarios is determined. This must be sufficient to reflect important variations in the analysis, but cannot include every possibility due to the huge number of possible failure scenarios. In this study, the number of failure scenarios is determined and samples generated as follows:

Assume c is the total number of components and c_f the number of failed components (i.e. the magnitude of stress). If only one component in the system fails ($c_f = 1$), this may be any one of the total set of c components and the total number of potential scenarios is c. When all components fail ($c_f = c$), there is only one possible scenario.

Given $1 < c_f < c$, two types of scenario are included in the sample development: *random failure* and *targeted failure*. Targeted failures are used to enrich the random sample and reduce the potential of missing critical scenarios, particularly those resulting indistinctive high and low strains. The *random failure scenario* picks all locations of the c_f failed components randomly, and each scenario is unique. Simulation of the *random failure scenarios* is repeated n_R times, where n_R is determined as follows (Brase and Brase, 2012):

$$n_R = \frac{N \cdot p \cdot (1-p)}{(N-1) \cdot D + p \cdot (1-p)} \tag{1}$$

$$D = \left(\frac{CI}{Z_\alpha}\right)^2 \tag{2}$$

here, n_R is the number of *random failure scenarios* modelled at each stress magnitude; N is the total number of possible scenarios for a failure mode; p is the probability of success. In this study, p is regarded as the probability of successful hydraulic simulation since hydraulic simulation may fail under a large stress magnitude (i.e. scenarios with failed simulation are invalid and the strains are not calculated). A value of $p = 0.5$ is assumed to provide the minimum sample size required. CI is the confidence interval; Z_α is the normal distribution value for a given confidence level (e.g. 1.645 at a confidence level of 90%, 1.960 at 95%, and 2.575 at 99%).

The *targeted failure scenarios* where $1 < c_f < c$ are built up incrementally by selecting c_f failed components one by one (Albert et al., 2000; Holme and Kim, 2002; He et al., 2009). Two different selection strategies are applied, resulting in two groups of *targeted failure scenarios* with an equal number of scenarios. One group is started by first selecting the individual component whose failure results in the greatest strain, and vice versa for the other group (i.e. select the component resulting in lowest). If there are multiple components causing the same level of strain, or no component causing any strain, one is randomly selected (He et al., 2009). At each subsequent selection (e.g. the c_f th), $(c_f - 1)$ failed components have been already selected, and only one more component (i.e. the c_fth failed component) is selected from remaining unselected components {i.e. $\left|c - (c_f - 1)\right|$}. Hence, to consider all possible locations of the c_fth failed component, the total number of *targeted failure scenarios* (n_I) should be $2\left|c - (c_f - 1)\right|$ at any given number of $(1 < c_f \leq c)$ failed components. As a complementary method to random failure, the targeted failure scenarios have been widely used in complex network science for exploring the resilience and/or robustness of the networks to failures at specific locations, e.g. target attack (Albert et al., 2000; Holme and Kim, 2002; He et al., 2009).

The total number of failure scenarios for evaluation is therefore $c + (n_R + n_I) + 1$, where c is the total number of scenarios with only one component failed ($c_f = 1$); $(n_R + n_I)$ the number of scenarios with more than one but less than all components failed $(1 < c_f < c)$; and 1 the number of scenarios with all components failed ($c_f = c$).

As an example, consider the resilience of a WDS with 428 pipes to the pipe failure mode (i.e. $c = 428$). The number of failure scenarios should be calculated for all possible magnitudes of stress, from a single pipe failure ($c_f = 1$) to failure of every pipe ($c_f = c = 428$). There are 428 ($=c$) possible single pipe failure scenarios, and one scenario when all pipes fail.

When $1 < c_f < c$, the number of scenarios is the sum of *random failure scenarios* (n_R) and *targeted failure scenarios* (n_I). The number of *random failure scenarios* (n_R) required is calculated as follows for $p = 0.5$, $CI = \pm 5\%$, and $Z_\alpha = 1.960$ for a confidence level of 95%:

$$n_R = \frac{\sum_{i=1}^{428} C_{428}^i \cdot (0.5) \cdot (1-0.5)}{\left(\sum_{i=1}^{428} C_{428}^i - 1\right) \cdot \left(\frac{0.05}{1.960}\right)^2 + (0.5) \cdot (1-0.5)} = 384$$

The number of *targeted failure scenarios* is $n_I = 2 \times \left|428 - (c_f - 1)\right|$, i.e. 854 for $c_f = 2$, 852 for $c_f = 3$, etc. Therefore, $(n_R + n_I) = 1242 - 2c_f$.

To sum up, the number of scenarios encompassing all stress magnitudes is:

$$\begin{cases} 428, & c_f = 1 \\ 1242 - 2c_f, & 1 < c_f < c \\ 1, & c_f = 1 \end{cases}$$

and the total number of scenarios is therefore $1671 - 2c_f$.

Step 5. Generate a resilience stress-strain curves showing the mean, maximum and minimum strains generated from the simulations for any given stress magnitude.

2.2. Failure modes selected

Failure modes in WDSs can be broadly categorised as either: (a) *a structural failure,* which results from malfunction of system components, or (b) *a functional failure,* where the emphasis is on system resilience under loading conditions beyond the design envelope (Mugume et al., 2015). In this work, three different failure modes are considered in order to evaluate WDS resilience from differing perspectives. These are pipe failure, excess demand, and substance intrusion. Responses to systematic pipe failure can reveal the resilience of the system to the loss of physical connectivity (structural). Responses to excess demand indicate resilience to additional point loads without structural failure (functional). Responses to substance intrusion reflect resilience of the system to water quality disturbance without a change of system structure or hydraulic loading. Hence, WDS resilience is comprehensively evaluated in terms of structure and function. For each case, the specific evaluation method is described. Investigating different failure modes individually helps us to distinguish the systems' dynamic response to specific failures. This is critical before moving to more complicated cases. Evaluation of multiple, simultaneous system failure modes (e.g. pipe failure and loss of power) is beyond the scope of this particular paper but a topic of further, ongoing research.

2.2.1. Pipe failure

In this failure mode, the stress is modelled by changing the status of pipes to completely closed for a duration of 3 h. This is an attempt to simulate the isolation of failed pipes (loss of connectivity) followed by repair within three hours (i.e. the typical no-penalty maximum response time for UK water service providers). The simulation is a simplification of a real pipe failure and does not include the process from occurrence of pipe failure (e.g. pipe burst) to detection and complete isolation. Thus, water loss before pipe isolation is not modelled. Failed pipes are considered as isolatable shortly after breakage, and hence the 'pipe failure' has been modelled by closing the pipes rather than opening them and allowing outflows to drain into the surroundings until tanks are emptied. The stress magnitude is the percentage of pipes failed, represented by red crosses in Fig. 1. Although pipe bursts typically occur during periods of low demand, pipe failure may result from a range of internal and external threats and can potentially occur at any time. In the simulation, pipe failure is, therefore, introduced to coincide with the peak demand period and so capture the maximum potential impact. The strain magnitude is assessed by calculating dD_f, the ratio of unsupplied demand to the total demand during strain duration (Eqs. (3) and (4)). The strain duration is the time between the first occurrence of supply failure and the final return to a non-failure mode. Time to strain is also measured, and is the time between the first application of the stress (i.e. start of pipe failure) and the first occurrence of supply failure resulting from the stress. For cases when a given magnitude of stress causes no strain, time to strain is not calculated.

$$q_i(t) = \begin{pmatrix} P_i(t) <= 0 : 0 \\ 0 < P_i(t) < P_{\min} : \quad d_i(t)[P_i(t)/P_{\min}]^{1/2} \\ P_i(t) >= P_{\min} : d_i(t) \end{pmatrix} \tag{3}$$

$$dD_f = \sum_{T_{SS}}^{T_{SE}} \sum_{i}^{N} \left(q_i(t) - q_{i,f}(t) \right) \Big/ \sum_{T_{SS}}^{T_{SE}} \sum_{i}^{N} q_i(t) \tag{4}$$

where $d_i(t)$— the expected nodal demand to supply at junction i at time t; $P_i(t)$— the actual pressure at junction i at time t; P_{min}— the minimum allowed pressure at junction i (20 m for nodes with demand, and 0 m for nodes without); $q_i(t)$— the estimated actual supplied nodal demand at junction i at time t when there is no pipe failure; $q_{i,f}(t)$— the estimated actual nodal demand at junction i at time t in failure scenario$f$; T_{SS}— start time of the strain; T_{SE}— end time of the strain; N— the total number of junctions.

2.2.2. Excess demand – fire fighting

This failure mode represents the effects of incrementally loading the WDS with excess nodal demands, above and beyond the nominal loading case. The stress is simulated by adding excess demands with specified magnitude and duration, such as those that might be experienced under fire fighting conditions, to nodes, (as illustrated by flames in Fig. 1). To enable comparison of the systems' resilience on a like-for-like basis, the stress is a 6 h duration fire flow of 26.67 L/s (ÖNORM B 2538, 2002) (equivalent to 422.73 US gallon/minute) for all systems, irrespective of the system's design criteria. The percentage of nodes subject to this additional demand represents the stress magnitude. The introduction of the demand is timed to coincide with the normal peak demand period. Whilst it is recognised that occurrence of fires may reduce normal demand, this is not guaranteed and the assumption of normal demand enables identification of critical scenarios. For example, a 100% stress magnitude does not necessarily mean that the whole city is ablaze: there may be several large fires which necessitate the use of every hydrant but normal demands still occur in unaffected areas.

The strain magnitude is assessed by calculating the ratio of the mean number of nodes with pressure deficiency at any time step during the strain duration to the total number of nodes. For nodes with demand, the pressure is regarded as insufficient if it is <17 m (ÖNORM B 2538, 2002) during fire fighting or <20 m (Marchi et al., 2014; Giustolisi et al., 2015) after fire fighting. For nodes without demand, the pressure is regarded as insufficient if it is ≤0 m (Marchi et al., 2014; Giustolisi et al., 2015). The strain duration is the time between first occurrence of insufficient node pressure resulting from the stress and the final recovery to sufficient pressure at all nodes. Time to strain is the time from when fire fighting starts to occurrence of first pressure deficiency caused by the stress.

2.2.3. Substance intrusion

This stress is simulated by imposing a 25 g substance mass

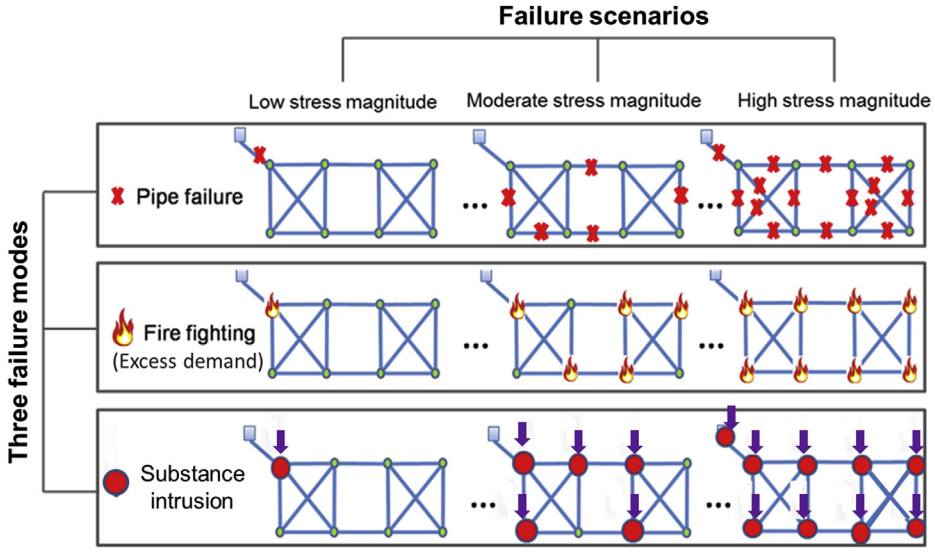

Fig. 1. Schematic of the three failure modes and corresponding failure scenarios. Crosses, flames and arrows represent the location of pipe failures, increased demands and substance intrusions respectively.

booster source at individual nodes for a duration of 2 h from the beginning of the simulation period (Zechman, 2011). Water supply is regarded as uncontaminated when a substance concentration is compliant with EU drinking water quality standards A threshold of <250 mg/L is used as applied to common substances such as chloride and sulphate [European Union (Drinking Water) Regulations, 2014]. The specific substance mass was chosen in order to 1) represent a substantial water quality deterioration violating the EU standard (i.e. 10 times the threshold value at source); 2) expose differences among systems and 3) avoid extensive computational load (high mass requires much longer simulation times as the system recovers much more slowly). The stress magnitude is modelled by the percentage of nodes to which an intrusion source is administered, represented by arrows in Fig. 1. The strain magnitude is evaluated by calculating the ratio of the volume of non-compliant water supply to the total water supply volume in strain duration. The strain duration is the time from first occurrence of non-compliant water supply until all water supplies in the system completely satisfies the threshold. In this study the intrusion is assumed to be undetected, and hence no emergency controls that may speed up recovery of the systems, are modelled. Hence, recovery of the systems depends solely on the system properties. Time to strain is the time from when substance intrusion starts to occurrence of contaminated supply.

A computer program was developed in C# to implement the GRA method for the three failure modes. The program generates failure mode scenarios by adding stresses in a WDS model through editing the model's input file. It then calls the EPANET engine (Rossman, 2000) for hydraulic and water quality simulation (demand-driven) of the scenarios. Based on simulation results, strain magnitude and duration are measured. Note that scenarios in which EPANET cannot solve the hydraulic equations (e.g. due to ill-conditioned equation systems) are omitted from the results plots. Analysis shows that these scenarios can be significantly different and do not result from failure of a group of critical elements. Hence, this is a weakness of the current numerical computation technique we used, and improving the fundamental hydraulic calculation scheme is out of the scope of this study. A pressure-driven model was also trialled but eventually discarded due to issues of convergence failure.

In the EPANET simulation, default settings are used (e.g. headloss formula, hydraulic time step) for hydraulic computation of all simulated WDSs. For the water quality simulation, the transport and decay of substances is modelled based on the principles of conservation of mass coupled with reaction kinetics (Rossman et al., 1993;Rossman and Boulos, 1996). The chosen settings for substance reaction in EPANET are: Bulk Reaction Order = 1.0; Global Bulk Coefficient = −1.0 (Rossman, 2000). Due to the high computational load of the water quality analysis, the maximum allowable water quality time step is used, i.e. the same as the demand pattern time step (e.g. 30 min or 1 h, depending on the model used).

3. Case studies

Four WDSs of different sizes and configurations were used as case studies and their properties and configurations are summarised in Fig. 2 and Table 1. The case study networks have been used as benchmarks for WDS studies on water quality (Ostfeld et al., 2008), calibration (Ostfeld et al., 2012), optimal design (Marchi et al., 2014; Giustolisi et al., 2015), and cascade failures (Sitzenfrei et al., 2011; Diao et al., 2014a). No modifications were made to the properties of the networks.

4. Results and discussion

This section illustrates the GRA results and discusses how they can contribute to the management and operation of water distribution systems. Each failure mode is first analysed individually and then the common features are explored.

The GRA results for the three failure modes are presented in Figs. 3–5, showing the magnitude of stress applied against the magnitude of strain on the system. The plots show the mean (dash dot), maximum (solid) and minimum (dash) relationships between stress and strain.

4.1. Pipe failure

In the case of pipe failure, the GRA curves (Fig. 3) indicate that the worst case WDS response to pipe failure is complete loss of water supply in all of the case study WDSs. Strain duration is either the same as the duration of pipe failure (i.e. 3 hrs) or slightly longer (e.g. 1 hr more). Time to strain (level of service failure) ranges from 0 h to 2 h. Thus, all four systems currently lack resilience to this failure mode, and the strain starts and stops immediately or shortly after pipe failure starts and ends, respectively. Specifically, the maximum strain curves of all systems (the solid lines in Fig. 3) rise sharply to reach the global maximum supply shortage (100%). Hence, failure of just a small fraction of pipes (approx. 6% for Net3; 5% for BWSN network 1; 29% for the Alpine network; and 25% for C-Town) can disrupt the systems' water supply entirely.

The results imply that, in each system, only a few pipes represent the most critical hydraulic links, e.g. trunks connecting reservoirs to the rest of the system. For instance, duplication of only 9 pipes [Fig. 4(A)] can dramatically increase the resilience of Alpine network to low stress yet high probability events [Fig. 4(B)]. These pipes are chosen via clustering analysis, as described by Diao et al. (2014a). Contrarily, failure of a large number of pipes does not necessarily result in catastrophic impacts. For instance, Net3 can still deliver 86% of total demand with 70% of pipes failed if critical pipes remain undamaged (shown in Fig. 3). The revealed criticality of pipes can guide investments in operation and maintenance and emergency planning (e.g. for pipe burst). Critical pipes are prioritized for protection. Uncritical pipes may be ideal sites for operation and maintenance control due to the low impact of disconnecting them. For example, those pipes can be closed to isolate a burst pipe nearby in emergency planning. Therefore, they can be ideal places to install isolation valves. As a result, the system can fail as planned to minimize the impacts. Based on the results, budget for increasing resilience to pipe failure can therefore be efficiently allocated.

Note that the analysis results may also be affected by the level of detail of the hydraulic models. Alpine and C-Town, for instance, reaches the maximum impact under a relatively large stress magnitude (29% and 25% respectively) in comparison with Net3 and BWSN network 1 (6% and 5%). This may be attributed partially to the greater detail in the Alpine and C-Town models. These two networks have a mean pipe length of 130 m (max 1600 m), compared with 562 m (max 3127 m) for Net3 and 224 m (max 13868 m) for BWSN Network 1. Consequently, failure of a pipe in Net3 or BWSN Network 1 may actually represent loss of an entire supply path. Loss of all connections in a trunk model does not necessarily mean failure of every individual pipe – instead, it represents loss of all paths to water sources due to failure of any pipes on those paths. Furthermore, increased model detail may lead to a larger number of possible scenarios, which in turn increases the difficulty of finding the most extreme scenarios. In this regard, it may be worthwhile starting the GRA with a trunk model, and then changing to detailed model (e.g. all-pipe model) if necessary (e.g. for detailed analysis of parts of the system found to be critical).

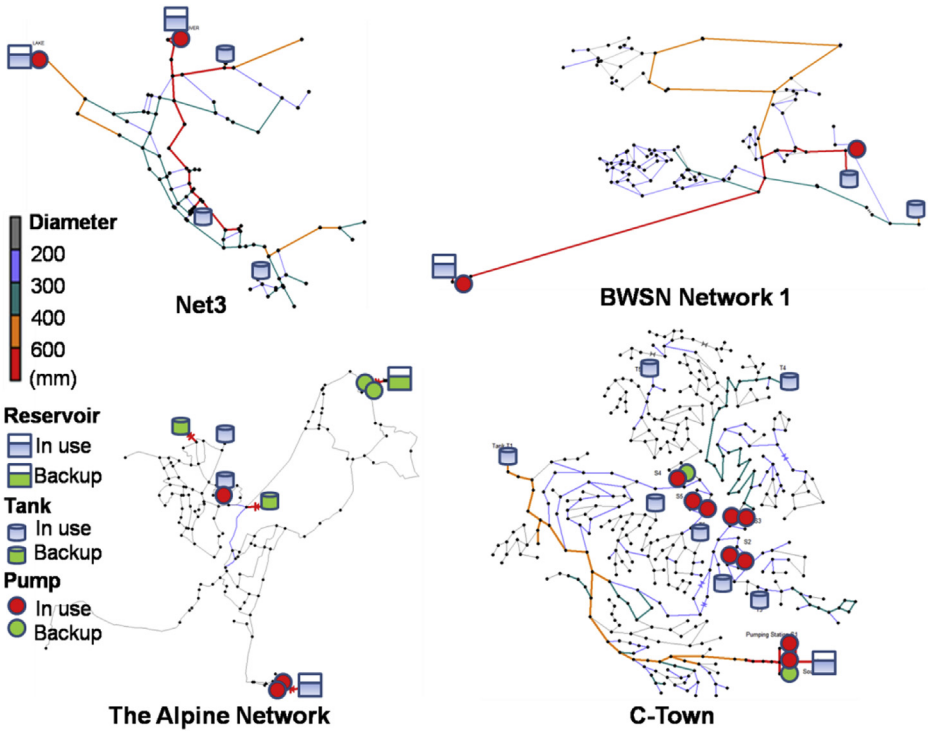

Fig. 2. Case study water distribution systems.

Table 1
General properties of the case study networks.

Networks	Net 3	Alpine	BWSN1	C-Town
No. of Nodes	91	127	126	388
No. of Links	115	157	168	429
No. of Reservoirs	2	2	1	1
No. of Tanks	3	4	2	7
No. of Pump stations	2	3	2	5
Simulation duration (hrs)	24	75	96	168
System mean demand (L/s)	717.3	6.36	56.37	170.16
Total tank volume (L)	28,621,000	732,358	35,148,000	9,496,639

4.2. *Excess demand—firefighting*

As shown in Fig. 5, the basic level of service (i.e. nodal pressures <17 m during firefighting and <20 m after) is breached for all the WDSs, although this occurs at various fractions of nodal coverage and has widely different strain durations starting at different times after the commencement of firefighting. Indeed, the systems each have very different response curves to these excess nodal demands.

The strain magnitude of Net3remains near zero (i.e. no or marginal level of service drop) until a large stress magnitude (number of fire fighting nodes) is imposed (i.e. 22% for maximum; 42% for mean; 68% for minimum), and so is considered highly resilient. Beyond this threshold, the lines then increase gradually and converge to a 65% loss of service when all nodes are stressed. The strain duration is no longer than 6 h, i.e. the duration of the fire fighting stress, and there is no strain when the magnitude of fire fighting stress is below 6%. Clearly, Net3 has a large buffer capacity to meet fire fighting demands that are relatively small with regard to the high demands of the system, e.g. system mean demand = 717.3 L/s. Quantifying the buffer capacity will also facilitate evaluation of the ability of WDSs to handle any demand uncertainties. With the exception of cases in which there is no strain (i.e. time to strain does not exist), the time to strain is either 2 or 0 h.

The BWSN network1 is also resilient to fire fighting stress, with a service drop threshold at 22%for max; 23% for mean; 55% for min. In contrast, the BWSN network1 has a much longer maximum strain duration (24 h) than Net3 (2 h), whilst it can absorb a much larger magnitude of fire fighting stress (16%) without any strain. Again, for scenarios with strains, the maximum time to strain of BWSN network1 is as long as 6 h, while that of Net3is only 2 h [Fig. 5] when the stress magnitude falls between 17% and 47%. However, strain in BWSN network1 begins earlier (maximum time to strain = 0 h) than in Net3 (maximum time to strain = 2 h) for a larger magnitude of fire fighting stress afterwards. Given these comparisons, it can be concluded that even for the same failure mode, a system can be more resilient than another with respect to one measure of strain and less resilient with respect to another. Even for the same measure of strain, a system can be more resilient than another under only a certain range of stress magnitudes and vice versa. These facts demonstrate the necessity of GRA, showing that analysis of a few scenarios which address only a limited range of potential stress magnitudes and types may not reveal the full dynamic of a system's response and tipping points.

Alpine and C-Town have low or even no buffer capacity to widespread excess demands, exhibiting rapid increases in strain over approximately the first 10% of stress. Correspondingly, the strain duration increases sharply to maxima (i.e. 7 and 17 h for Alpine and C-Town respectively) at only 2% (Alpine) and 5% (C-Town) node coverage. Nevertheless, the time to strain of the two systems has a large range as well, although they cannot absorb any magnitude of fire fighting stress with no strain. For example, in some cases the strain on C-Town does not occur until 83 h (Fig. 4) after the stress was imposed.

The exceptionally long strain durations in BWSN and C-Town are attributed to the long recovery time of only a few nodes. The long-lasting pressure deficiencies at those nodes result from drained

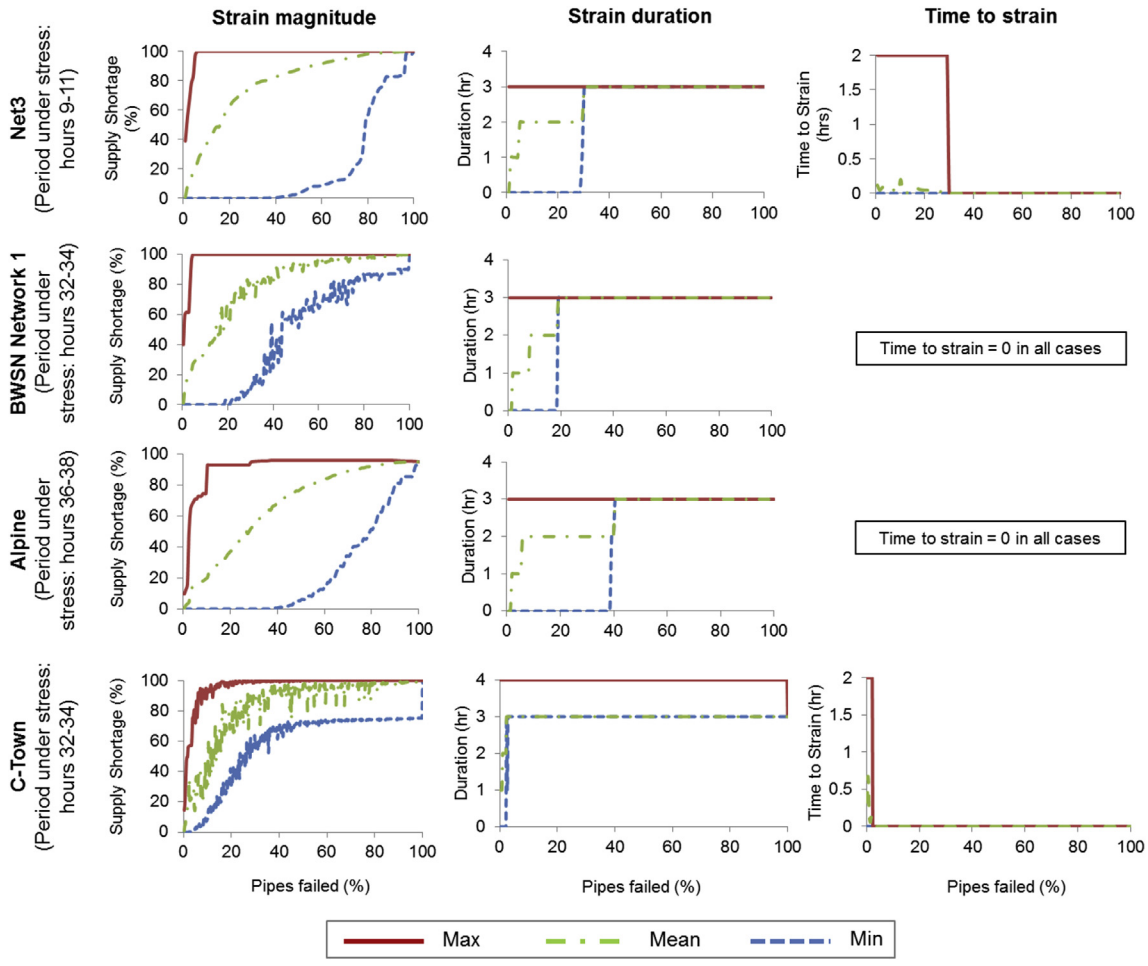

Fig. 3. GRA curves for pipe failure.

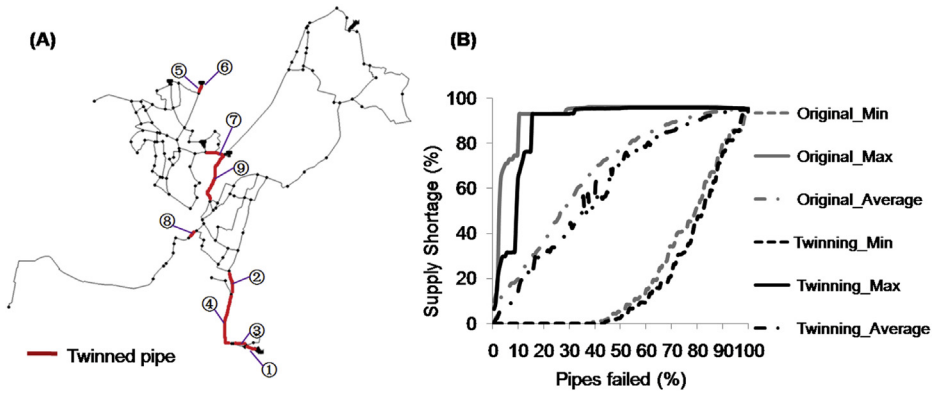

Fig. 4. Enhancement of the Alpine network resilience to pipe failure by twinning critical pipes.

tanks and lack of pumping capacity (e.g. to refill the tanks quickly). This is understandable as the WDSs are not sized to cope with exceptional high impact failures. However, this also implies that increasing capacity may not necessarily enhance resilience, since dramatically increased capacity (e.g. tank size) may still be too small to reduce impacts of high stress, and worse prolong the recovery time significantly after removal of the stress (e.g. due to the longer time taken to refill large tanks). Furthermore, increased hydraulic capacity may have negative effects on water quality and this should be fully explored before implementing such interventions to increase resilience to excess demand. Hence, increase of capacity should be very carefully planned. To shorten recovery time, more operational options (e.g. more available pumps to increase pumping capacity) would be more feasible and cost-efficient as emergent post-stress responses. For instance, by adding one pump at the water source in BWSN Network 1 (with the same property as the original one, PUMP-172) [Fig. 6(A)] and running it for 5 h (hours 38–43) after fire-fighting, the maximum

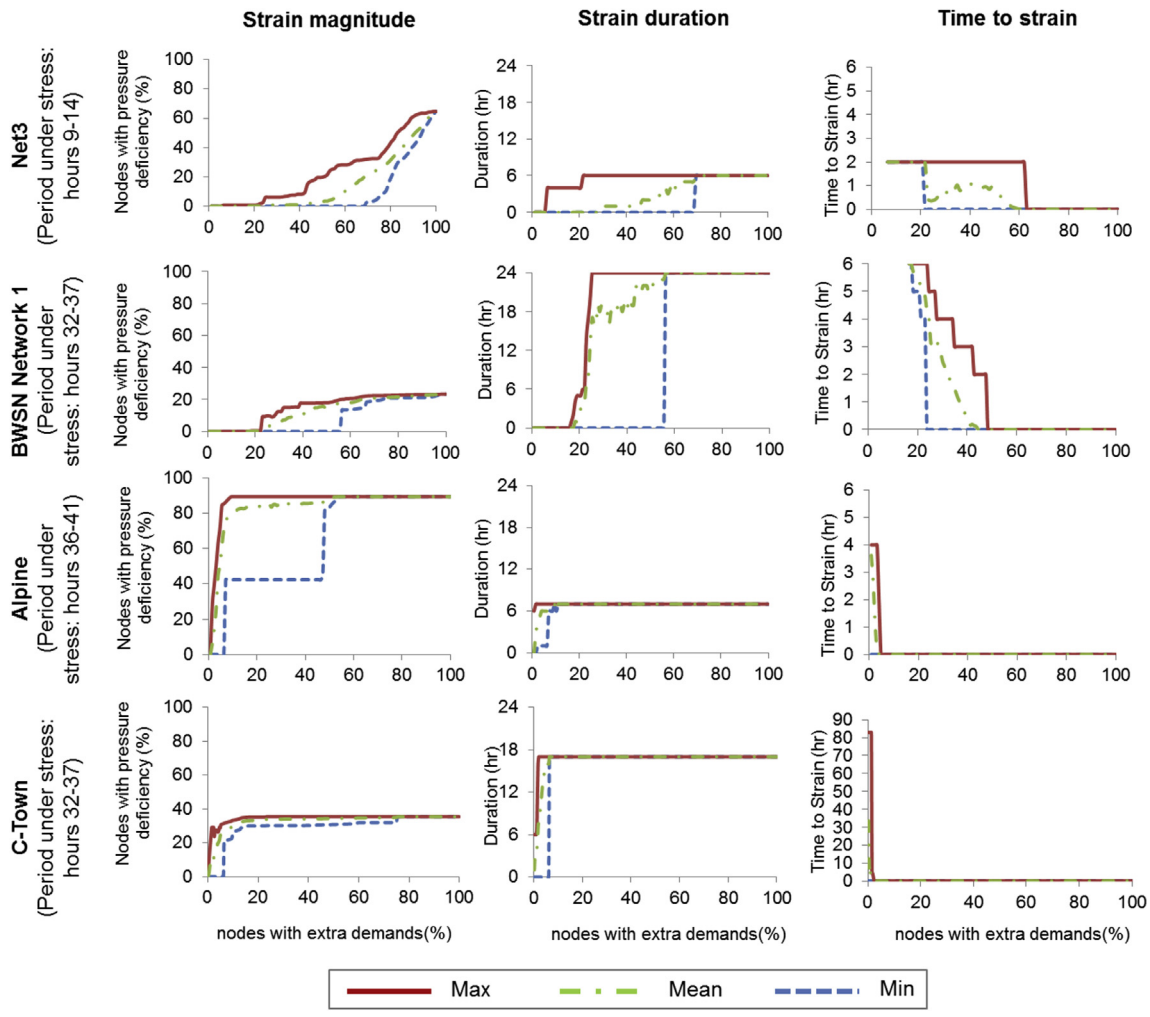

Fig. 5. GRA curves for excess demand (firefighting).

strain duration is reduced by 12 h [Fig. 6(B)]. This indicates that a guideline for resilient design of WDSs is definitely required, since such an extra pump would not be regarded as necessary if failure scenarios were not considered. This also demonstrates that assessing resilience enhancement due to implementation of different interventions should include properties and performance.

For this failure mode, the level of model detail may also affect the analysis and results. In less-detailed models, many locations of hydrants maybe missing and therefore use of multiple nearby hydrants for fire-fighting cannot be modelled. In all pipe models, nodes configured with hydrants should be identified as it may not appropriate to assume a hydrant is available at every node. Hence, the appropriate level of model detail is an area for future exploration, and knowledge of all hydrant locations is a prerequisite.

4.3. Substance intrusion

Fig. 7 shows GRA curves for the four networks in response to substance intrusion events. For this failure mode, time to strain is regarded as 0 for all cases, as substance intrusion starts at the very beginning of the model simulation. Hence, it is not included in Fig. 7.

There is variety in these curves indicating resilience to substance intrusion is very system specific. For instance, the maximum strain magnitude (i.e. percentage of contaminated water supply) ranges from 1.5% (Net3) to 100% (Alpine). The maximum strain duration ranges from 3 h (Net3) to 140 h (C-Town). Net3 is the most resilient to substance intrusion, followed by BWSN network1. The high flow rates due to large water demands in these two networks dilute the concentration of substance. As for the other two, C-Town has a lower maximum strain magnitude (57%) than Alpine (100%), but a longer maximum strain duration (140 h compared with 42 h). The large variation in strain durations may attributed to different reasons. For instance, the long strain duration of C-Town may be explained by the wide spread of tanks over the network, each of which needs a long time to recover once contaminated. Accordingly, the maximum strain duration of C-Town can be shortened to 100 h by isolation of all tanks from the beginning of the intrusion until the system recovers. To meet the demands, however, the backup pump (PU3) at the water source needs to be switched on until the system recovers.

As for water quality monitoring and control, the results provide some clues for sensor placement. For instance, the maximum impact in Alpine increases rapidly and exceeds80% contamination when the stress magnitude is only 6%. This indicates that a few nodes once polluted will have broad impact on the system. Hence, those nodes can be good candidates for sensor placement. Contrarily, the maximum impact in BWSN Network 1 and C-Town increases gradually, which is an indicator of good resilience. However, it reveals no node is significantly more important than any

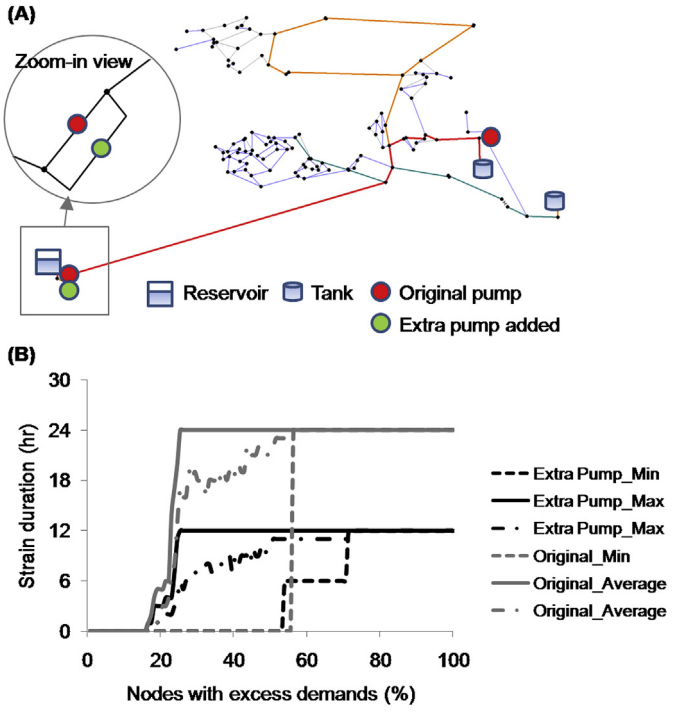

Fig. 6. Shortened recovery of the BWSN following excess demand by increasing pumping capacity.

other for the studied substance intrusion. Consequently, more analysis (e.g. intrusion of different substances) may have to be done to identify suitable sites for sensors.

Regarding the level of model detail, models containing all pipes are required for accurate evaluation of resilience to substance intrusions, as it is unrealistic to model pollution of a pipeline several thousand meters long (e.g. as in BWSN network 1) becoming polluted in one time-step. However, water quality simulation for such a detailed model could be very time consuming. A trade-off between accuracy and efficiency may be necessary, and therefore requires further future research.

4.4. Common features

For all the three failure modes, there were some common features:

The range of strain magnitude first increased as stress magnitude increased and then reduced after the stress magnitude passed a threshold (different from system to system and failure mode to failure mode), finally converged when the whole WDSs were under stress.

The same magnitude of stress could result in very different level of strains, as shown by the gap between maximum and minimum curves. For instance, the strain magnitude resulting from single pipe failure varies from 0% to about 40% in Net3 and BWSN network1, 0–15% in C-Town, and 0–10% in Alpine. Moreover, it ranges from nearly 0% to 100% when the fraction of pipes failed reaches 40% in Net3 and Alpine and 20% in BWSN network1; and from nearly 9% to 100% in C-Town with 10% of pipes failed. As discussed before, these big gaps reveal the existence of critical scenarios and the different levels of model detail. For instance, when 70% of pipes fail Net3 can be completely out of service (100% supply shortage) or still deliver about 86% of total demand, depending on the location of the failed pipes. The latter scenario meets the majority of demands with a tremendously high proportion of pipes failed because there is no pipe failure in the backbone (Diao et al., 2014b, 2015) formed by the most critical links connecting critical infrastructures (e.g. reservoirs, tanks, pump stations, large customers). Hence, enhancing the resilience of backbones to failure is crucial. For the other two failure modes, different combinations of failed components can lead to considerably varied strains as well. For the excess demand, the strain magnitude can range from 3.3% to 32% at a stress magnitude of 74% in Net3, and from 0% to 19% at a stress magnitude of 56% in BWSN network1, and from 46% to 90% at a stress magnitude of 47% in Alpine. For the substance intrusion, the range of strain magnitude at a stress magnitude of 70% varies from 3.4% to 21.6% for BWSN network 1, from 53% to 100% for Alpine and from 13.5% to 57% for C-Town.

Increased resilience to one failure mode may decrease resilience for another. For instance, although tanks enhance flexibility and capacity to meet demand [e.g. as temporary water sources when the reservoir(s) is disconnected], they could be a negative factor regarding resilience to substance intrusion. This is demonstrated in C-Town, where the wide spread of several tanks can either prevent supply shortage resulting from pipe failure (i.e. give zero strain) or delay the occurrence of supply shortage (e.g. time to strain = 2, Fig. 3). Contrarily, once those tanks are contaminated, the sparse spatial distribution of them may facilitate diffusion of contaminated water within the system, and results poor resilience to contaminant intrusion (as illustrated by the 140 h recovery time in Fig. 5).

Additionally, given that this study investigates resilience to specific failure modes, it is expected that the results may differ from those obtained using other resilience indices. Whilst a comparison would be interesting, it is beyond the scope of this work.

5. Conclusions

This paper proposes a new method, global resilience analysis (GRA), and applies it to water distribution systems (WDSs). The GRA described evaluates the resilience off our benchmark systems to three different failure modes, those being pipe failure, excess demand (e.g. firefighting), and substance intrusion. A failure mode includes all scenarios of the system under a particular stress, regardless of what threat(s) causes the stress. The GRA results are presented in a series of plots which show the relationship between stress magnitude and multiple indicators of strain. From examination of the results of the GRA for each failure mode and WDS, it can be concluded that:

1) GRA can identify: (i) the level of resilience of the same system to different failure modes; (ii) the level of resilience of different systems to the same failure mode; (iii) the range of strains (minimum and maximum) which may result from any given level of stress; (iv) the stress magnitude the system can withstand before reaching a certain level of service reduction; (v) the existence of thresholds where a slight increase in stress gives rise to more severe impacts; and (vi) scenarios which result in the minimum and maximum strains, which can reveal combinations of components whose failure would lead to marginal or considerable strains. Therefore, this method can be used as comprehensive diagnostic framework linking system attributes (e.g. connectivity and capacity) to performance (e.g. level of service) to evaluate how best to build whole-system resilience based on multiple system failure modes in future studies.
2) The GRA illustrates the similarities or differences in complex dynamic responses of different systems to various failure modes. For different failure modes, a system can be more resilient to one failure mode than another system while be less

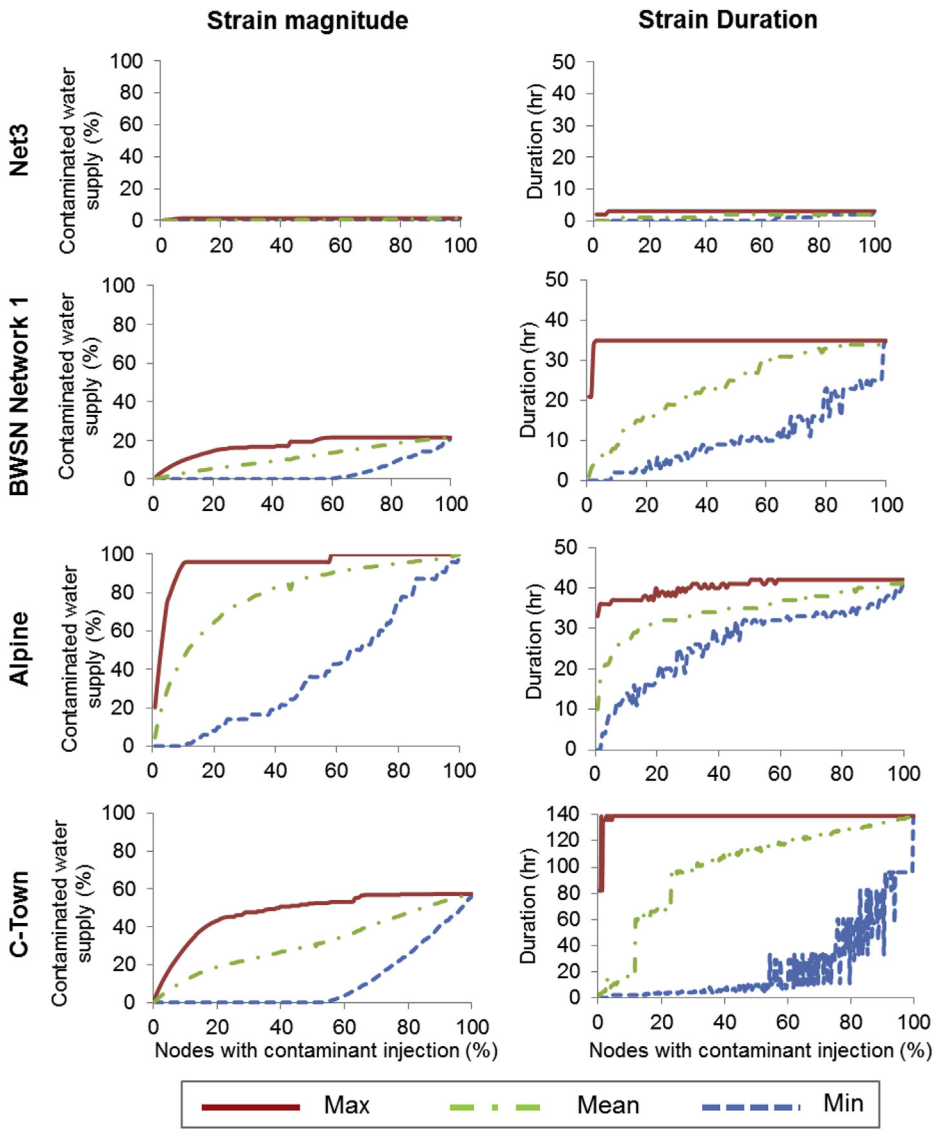

Fig. 7. GRA curves for substance intrusion.

resilient to another failure mode. For the same failure mode, a system can be more resilient than another system with respect to one measure of strain and less resilient with respect to another. Even for the same measure of strain, a system can be more resilient than another under a certain range of stress magnitude and vice versa.

3) The GRA results reveal that increased resilience to one failure mode may decrease resilience to another. For example, tanks enhance flexibility and capacity to meet demands [e.g. as temporary water sources when the reservoir(s) is disconnected], yet they could be a negative factor regarding resilience to substance intrusion (e.g. take very long time to recover once contaminated).
4) Extreme scenarios identified (e.g. those with very high or low impacts) can guide resilience enhancement, and be used to ensure systems fail as planned during unexpected stresses. For instance, in some scenarios failure of a small number of critical components (e.g. 5%) may disable the entire system for water supply, whilst in other scenarios 80% of demand may still be met when 70% of components have failed. Hence, different operation and maintenance strategies can be tailor-made for critical and uncritical components.
5) Increasing capacity (e.g. by increasing tank size) may not always improve resilience, and may even prolong the system's recovery time significantly following removal of the stress. It is therefore necessary to assess both the properties and performance of any intervention designed to improve resilience.
6) The level of detail of hydraulic models may affect the GRA results, and therefore the level of model detail required should be further explored for different failure modes.

Acknowledgements

The work reported is funded by the UK Engineering & Physical Sciences Research Council (EPSRC) project Safe & SuRe (EP/K006924/1).

References

Albert, R., Jeong, H., Barabási, A.L., 2000. Error and attack tolerance of complex

networks. Nature 406, 378–382.
Ahern, J., 2011. From fail-safe to safe-to-fail: sustainability and resilience in the new urban world. Landscape 100, 341–343.
ASCE Policy Statement 518. (2006) http://www.asce.org/issues-and-advocacy/public-policy/policy-statement-518--unified-definitions-for-critical-infrastructure-resilience/ (Accessed 10 Febraury 2015).
Berardi, L., Ugarelli, R., Røstum, J., Giustolisi, O., 2014. Assessing mechanical vulnerability in water distribution networks under multiple failures. Water Resour. Res. 50, 2586–2599. http://dx.doi.org/10.1002/2013WR014770.
Bristow, E., Brumbelow, K., Kanta, L., 2007. Vulnerability assessment and mitigation methods for interdependent water distribution and urban fire response systems. In: World Environmental and Water Resources Congress, pp. 1–10.
Butler, D., Farmani, R., Fu, G., Ward, S., Diao, K., Astaraie-Imani, M., 2014. A new approach to urban water management: safe and SuRe. Procedia Eng. 89, 347–354.
Butler, D., Ward, S., Sweetapple, C., Astaraie-Imani, M., Diao, K., Farmani, R., Fu, G., 2016. Reliable, resilient and sustainable water management: the Safe &SuRe approach. Global Chall. http://dx.doi.org/10.1002/gch2.1010.
Brase, C., Brase, C., 2012. Chapter 7: Estimation in Understandable Statistics: Concepts and Methods. Thomson Brooks/Cole, 10 Pck Pap edition (6 Jan. 2012).
Cabinet Office, 2011. Keeping the Country Running: Natural Hazards and Infrastructure. A Guide to Improving the Resilience of Critical Infrastructure and Essential Services. Civil Contingencies Secretariat, Cabinet Office.
Diao, K.G., Farmani, R., Fu, G.T., Astaraie-Imani, M., Ward, S., Butler, D., 2014a. Clustering analysis of water distribution systems: identifying critical components and community impacts. Water Sci. Technol. 70 (11), 1764–1773.
Diao, K.G., Fu, G.T., Farmani, R., Guidolin, M., Butler, D., 2014b. Hierarchical decomposition of water distribution systems for background leakage assessment. Procedia Eng. 89, 53–58.
Diao, K.G., Fu, G.T., Farmani, R., Guidolin, M., Butler, D., 2015. Twin hierarchy decomposition for optimal design of water distribution systems. J. Water Resour. Plan. Manag. (Special Issue on The Battle of Background Leakage Assessment for Water Networks, accepted).
European Union (Drinking Water) Regulations 2014. S.I. No. 122 of 2014. The Stationery Office, Dublin. Available from: https://www.fsai.ie/uploadedFiles/Legislation/Food_Legisation_Links/Water/SI122_2014.pdf.
Francis, R., Bekera, B., 2014. A metric and frameworks for resilience analysis of engineered and infrastructure systems. Reliab. Eng. Syst. Saf. 121, 90–103.
Gheisi, A., Naser, G., 2014. Water distribution system reliability under simultaneous multicomponent failure scenario. J. Am. Water Works Assoc. 106 (7), E319–E326.
Giustolisi, O., Berardi, L., Laucelli, D., Savic, D., Kapelan, Z., 2015. Operational and tactical management of water and energy resources in pressurized systems: the competition at WDSA 2014. J. Water Resour. Plan. Manag. http://dx.doi.org/10.1061/(ASCE)WR.1943-5452.0000583 (Special Issue on The Battle of Background Leakage Assessment for Water Networks).
He, S., Li, S., Ma, H., 2009. Effect of edge removal on topological and functional robustness of complex networks. Phys. A 388 (2009), 2243–2253.
Hokstad, Per, Utne, Ingrid B., Vatn, Jørn, 2013. Risk and Interdependencies in Critical Infrastructures. Chapters 5 and 6. Springer, pp. 49–94.
Holme, P., Kim, B., 2002. Attack vulnerability of complex networks. Phys. Rev. E 65, 056109–1–14.
Hughes, J.F., Healy, K., 2014. Measuring the Resilience of Transport Infrastructure. NZ Transport Agency research report 546. AECOM New Zealand Ltd.
Johansson, J., 2007. Risk and Vulnerability Analysis of Large-scale Technical Infrastructures: Electrical Distribution Systems (Doctoral dissertation). Lund University, Sweden.
Johansson, J., Henrik, H., 2010. An approach for modelling interdependent infrastructures in the context of vulnerability analysis. Reliab. Eng. Syst. Saf. 95, 1335–1344.
Johansson, J., Henrik, H., Zio, E., 2013. Reliability and vulnerability analyses of critical infrastructures: comparing two approaches in the context of power systems. Reliab. Eng. Syst. Saf. 120, 27–38.
Kanta, L., Brumbelow, K., 2013. Vulnerability, risk, and mitigation assessment of water distribution systems for insufficient fire flows. J. Water Resour. Plan. Manag. 139 (6), 593–603.
Kanta, L., 2006. Vulnerability Assessment of Water Supply Systems for Insufficient Fire Flows (Master thesis). Texas A&M University.
Marchi, A., Salomons, E., Ostfeld, A., Kapelan, Z., Simpson, A., Zecchin, A., Maier, H., Wu, Z., Elsayed, S., Song, Y., Walski, T., Stokes, C., Wu, W., Dandy, G., Alvisi, S., Creaco, E., Franchini, M., Saldarriaga, J., Páez, D., Hernández, D., Bohórquez, J., Bent, R., Coffrin, C., Judi, D., McPherson, T., van Hentenryck, P., Matos, J., Monteiro, A., Matias, N., Yoo, D., Lee, H., Kim, J., Iglesias-Rey, P., Martínez-Solano, F., Mora-Meliá, D., Ribelles-Aguilar, J., Guidolin, M., Fu, G., Reed, P., Wang, Q., Liu, H., McClymont, K., Johns, M., Keedwell, E., Kandiah, V., Jasper, M., Drake, K., Shafiee, E., Barandouzi, M., Berglund, A., Brill, D., Mahinthakumar, G., Ranjithan, R., Zechman, E., Morley, M., Tricarico, C., de Marinis, G., Tolson, B., Khedr, A., Asadzadeh, M., 2014. Battle of the water networks II. J. Water Resour. Plan. Manag. 140 (7), 04014009.
Mugume, S.N., Gomez, D.E., Fu, G., Farmani, R., Butler, D., 2015. A global analysis approach for investigating structural resilience in urban drainage systems. Water Res. http://dx.doi.org/10.1016/j.watres.2015.05.030.
NIAC, 2009. Critical Infrastructure Resilience - Final Report and Recommendations. US National Infrastructure Advisory Council.
Ofwat, 2012. Resilience – Outcomes Focused Regulation. Principles for Resilience Planning. The Water Services Regulation Authority for England &Wales.
ÖNORM B 2538, 2002. Long-distance, District and Supply Pipelines of Water Supply Systems – Additional Specifications Concerning ÖNORM EN 805. Österreichisches Normungsinstitut, Wien.
Ostfeld, A., Uber, J., Salomons, E., Berry, J., Hart, W., Phillips, C., Watson, J., Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F., Khu, S., Savic, D., Eliades, D., Polycarpou, M., Ghimire, S., Barkdoll, B., Gueli, R., Huang, J., McBean, E., James, W., Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., VanBriesen, J., Small, M., Fischbeck, P., Preis, A., Propato, M., Piller, O., Trachtman, G., Wu, Z., Walski, T., 2008. The battle of the water sensor networks (BWSN): a design challenge for engineers and algorithms. J. Water Resour. Plan. Manag 134 (6), 556–568.
Ostfeld, A., Salomons, E., Ormsbee, L., Uber, J., Bros, C., Kalungi, P., Burd, R., Zazula-Coetzee, B., Belrain, T., Kang, D., Lansey, K., Shen, H., McBean, E., Wu, Z., Walski, T., Alvisi, S., Franchini, M., Johnson, J.P., Ghimire, S.R., Barkdoll, B.D., Koppel, T., Vassiljev, A., Kim, J.H., Chung, G., Yoo, D.G., Diao, K.G., Zhou, Y.W., Li, J., Liu, Z.L., Chang, K., Gao, J., Qu, S., Yuan, Y., Prasad, T.D., Laucelli, D., VamvakeridouLyroudia, L.S., Kapelan, Z., Savic, D., Berardi, L., Barbaro, G., Giustolisi, O., Asadzadeh, M., Tolson, B.A., McKillop, R., 2012. The battle of the water calibration networks (BWCN). J. Water Resour. Plan. Manag. ASCE 138 (5), 523–532.
Pickett, S.T.A., McGrath, B., Cadenasso, M.L., Felson, A.J., 2014. Ecological resilience and resilient cities. Build. Res. Inf. 42 (2), 143–157.
Rossman, L.A., 2000. EPANET 2 Users Manual. U.S. Environmental Protection Agency, Cincinnati.
Rossman, L., Boulos, P., 1996. Numerical methods for modeling water quality in distribution systems: a comparison. J. Water Resour. Plan. Manag. 122 (2), 137–146.
Rossman, L., Boulos, P., Altman, T., 1993. Discrete volume-element method for network water-quality models. J. Water Resour. Plng. Mgmt. 119 (5), 505–517.
Sitzenfrei, R., Mair, M., Möderl, M., Rauch, W., 2011. Cascade vulnerability for risk analysis of water infrastructure. Water Sci. Technol. 64 (9), 1885–1891.
Zechman, M., 2011. Agent-based modeling to simulate contamination events and evaluate threat management strategies in water distribution systems. Risk Anal. 31 (5), 759–772.